

Elektronische Nutzungsstatistiken – oder: Zu Fuß ist es kürzer als über den Berg

Miriam Lorenz

This document appeared in

Bernhard Mittermaier (Eds.):

eLibrary - den Wandel gestalten

5. Konferenz der Zentralbibliothek

Proceedings of the WissKom 2010: 5. Konferenz der Zentralbibliothek, 08.-11. November 2010, Jülich

Schriften des Forschungszentrums Jülich / Reihe Bibliothek/Library, Vol. 20

Zentralbibliothek (ZB)

Forschungszentrum Jülich GmbH, Zentralbibliothek, Verlag, 2010

ISBN: 978-3-89336-668-2

Elektronische Nutzungsstatistiken – oder: Zu Fuß ist es kürzer als über den Berg

Miriam Lorenz

Zusammenfassung

Die Bedeutung der Nutzungsstatistiken von elektronischen Ressourcen scheint unbestritten. Mit Hilfe dieser Daten kann der Nutzen und damit die Begründung für die Lizenzierung nachgewiesen werden – oder? Dieser Übersichtsartikel bietet, beruhend auf Erfahrungswissen, Austausch und der Analyse des aktuellen Forschungsstandes erstmalig eine Zusammenfassung über Probleme, Schwierigkeiten und Fehler in Zusammenhang mit Nutzungsstatistiken. Die aufgezeigten Mängel sind in allen Schritten bei der Erstellung, Verarbeitung und Interpretation von Nutzungsstatistiken zu finden: z.B. technische und organisatorische Limitationen, alte Konzepte zur Identifizierung von Zeitschriften oder auch die Einschränkung in der Umsetzung von Nutzungsdaten durch Vertragsmodelle wie den „Big Deal“¹. Abschließend wird festgehalten, dass durch Fehlern auf verschiedenen Ebenen eine Vergleichbarkeit der Statistiken nicht gegeben ist und eine Aussage über den Nutzen oder eine Lizenzentscheidung nicht auf Grundlage von Nutzungsstatistiken getroffen werden kann. Jedoch können die Statistiken Hinweise auf das Nutzungsverhalten im Kontext von elektronischen Inhalten geben und eine gute Grundlage für die Rechtfertigung von Budgets bieten. Weiter werden die Punkte aufgezeigt, durch welche Anbietern und Informationseinrichtungen sich weiter dem Ideal einer verlässlichen Nutzungsstatistik nähern können.

Abstract

The importance of "user statistics" of electronic resources seems to be not controversial. Using these dates you are enabled to prove the advantage and though the explanatory statement of licensing – aren't you? This is a survey primarily providing an abstract of problems, difficulties and failures with regard to user statistics. The abstract is based on know-how, exchange of information and analysis of the current state of research. The deficits are above all to be found when creating, processing and interpreting user statistics e. g. technical or organizational limitations, antiquated concepts for identification of journals or the limitation of the adoption of user dates

¹ d.h. dem Abonnement aller elektronischer Zeitschriften eines Verlages ohne jede Auswahl

e. g. concepts of contracts like "Big Deal"². In conclusion I state that based on mistakes on various levels a comparability of statistics is not possible. A statement relating to the use or the decision of licensing can't be made. However statistics allow you to get hints to usage behaviour of electronic contents though offering a good basis for the justification of budgets. Furthermore possibilities are shown enabling providers and information institutes to approach the ideal of a reliable user statistic.

Einführung

In einer Befragung der Mitglieder der Association of Research Libraries (ARL) zum Thema „Anbieterstatistiken elektronischer Ressourcen“ im Jahr 2008 gaben 94% der Bibliotheken an, Nutzungsstatistiken für „subscription decisions“ zu verwenden (Baker/Read 2008). Mit dieser Aussage beschreiben die Bibliotheken wohl weniger eine dezidiert auf Nutzungsverhalten gründende Erwerbungspolitik, sondern – in Ermangelung von Alternativen - mehr die Berücksichtigung von groben Hinweisen oder Tendenzen. Denn derzeit führt die Analyse der Erstellung, Aufbereitung und Vergleichbarkeit von Statistiken elektronischer Quellen, ebenso wie Anwendbarkeit möglicher Interpretation der Statistiken zu einem bunten „Problem-Potpourri“. Dieser besteht - wie später näher erläutert - aus technischen Limitationen, Fehlern, fehlenden Standards, ungenauen Definitionen, enormen Zeitaufwand und unflexiblen Strukturen auf Seiten aller Beteiligten. Doch trotz dieser zum Teil enormen Defizite der Statistiken scheint das Interesse und die Verwendung der Daten in den Bibliotheken aller Typen weltweit zu steigen. Es finden nationale und internationale Workshops statt, schnelle Hilfe und Antworten erhält man auf fachspezifischen Mailinglisten (z.B. lib-stats) und insbesondere in amerikanischen und englischen Fachzeitschriften wird der Einsatz von Statistiken diskutiert und beschrieben - das „Journal of Electronic Resources Librarianship“ widmete diesem Themenkomplex 2009 gar ein Sonderheft (McDonald 2009). Gleichzeitig wird in der Standardisierungsinitiative COUNTER³ (u.a. Shepherd 2007) oder dem Forschungsprojekt MESUR⁴ (u.a. Bollen/Rodriguez/Van de Sompel 2007) u.a. versucht den Statistiken durch die Entwicklung eines sogenannten Usage Factors als Konkurrenz für den derzeit verwendeten Impact Factor die zentrale Stellung in der Wissenschaft und durch die Bildung einer Alternative zum Impact Factor auch auf dem Informationsmarkt zuzuweisen.

² That means the subscription of any journals of a publishing company without any selection.

³ <http://www.projectcounter.org>

⁴ <http://www.mesur.org>

Auch die Fachinformationsverantwortlichen der Fraunhofer-Gesellschaft haben das Potential von Statistiken erkannt und Kapazitäten für die Analyse von Statistiken bereitgestellt. Doch blieb auch hier die Frage offen: sind die derzeitig zur Verfügung stehenden Statistiken von elektronischen Ressourcen ein unbrauchbarer Zahlensalat oder die wichtigste Grundlage elektronischer Lizenzierung und revolutionär für den Zeitschriftenmarkt?

Im Folgenden wird durch die Gegenüberstellung der Probleme und Möglichkeiten von elektronischen Nutzungsstatistiken als Kombination aus Erfahrungswissen, Austausch und die Auswertung von Artikeln, insbesondere aus dem englisch-sprachigen Ausland, versucht, diese Frage zu beantworten.

Ein Hoch auf COUNTER

In ungefähr zehn Jahren Statistiken von elektronischen Ressourcen hat sich viel verändert. Vor allem die Standards, welche im Projekt COUNTER in Kooperation von Anbietern und Bibliotheken entstanden sind, stellen einen Meilenstein dar. Im sogenannten „COUNTER Code of Practice“ werden u.a. der genaue Aufbau der Tabellen, die Definition einer Zahl bzw. eines Begriffes oder auch das Dateiformat und die Art der Lieferung der Tabelle festgelegt und ständig weiterentwickelt. Kaum ein Artikel zum Thema Nutzung bzw. Statistik, in dem die Beschreibung von COUNTER nicht mindestens einen Absatz oder gar ganze Seiten füllt. Diese Fokussierung auf COUNTER spiegelt die Begeisterung der Bibliotheken wieder: „A number of the respondents praised the efforts of COUNTER and acknowledged that the COUNTER Code of Practice has been helpful in facilitating the use of vendor data...“ (Baker/Read 2008: S. 55). Artikel und Studien aus der „Vor-COUNTER“-Zeit beschäftigen sich fast ausschließlich mit der Notwendigkeit von Standards und empfehlen eine Beschäftigung mit Statistiken nur in Kombination mit großem Expertenwissen (Peters 2002). Seit der Einführung von COUNTER-kompatiblen Statistiken wird zunehmend die Anwendung und der große Nutzen und Erkenntnisgewinn von Statistiken auch im Vergleich zu den Daten über Print-Medien betont (Taylor-Roe/Spencer 2005; Moore/MacCreery/Marlow 2010).

COUNTER stellt damit sicher die wichtigste Entwicklung und die Lösung für viele anfängliche Probleme dar – und ist dabei doch nur ein Tropfen auf den heißen Stein.

Nutzungsstatistik = Statistik über das, was nutzt??

Statistiken von elektronischen Ressourcen werden nahezu ausschließlich in einem Atemzug mit „Nutzung“ genannt. Denn sie werden erhoben, um die Nutzung der lizenzierten Produkte zu zeigen. Jedoch steht Nutzung in keinem Zusammenhang mit dem Nutzen für den Empfänger. (Best 2007). Insbesondere zeigt dies die wohl meist verbreitete „Nutzungszahl“ im Bereich der eJournals: Artikel-Download. EIN Download wird hier gleichgesetzt mit EINER Nutzung. Ob dieser Artikel dabei gelesen, gedruckt und/oder gespeichert wurde, ist völlig unklar. Möglich ist auch, dass der Artikel von dem Nutzer immer wieder neu bei Bedarf runtergeladen wird und er somit immer wieder neu „Nutzung“ erzeugt. Und ob dieser Artikel dann letztendlich dem Leser etwas „genutzt“ hat – bleibt ebenfalls unbekannt. Selbstverständlich haben Nutzungsstatistiken elektronischer Medien dies mit den bewährten Zahlen zur Erfassung von Print-Medien (z.B. Ausleihe) gemeinsam. Doch führt diese Unklarheit über den Nutzen von elektronischen Medien zu bisweilen ungewöhnlichen Interpretation wie z.B. dass 0-Nutzung nicht „kein Nutzen“, sondern ein Potential an Nutzen beinhaltet⁵. Dies würde in der Umsetzung bedeuten, dass unabhängig von jeder Download-Zahl eine Zeitschrift Nutzen haben kann und macht damit letztendlich jede Statistik überflüssig. Doch selbst wenn man diese Behauptung ablehnt bleibt die Frage: Steht eine höhere Nutzung für mehr Nutzen? Zur Klärung wird im bibliothekarischen Alltag in der Regel ergänzend die direkte Bewertung des Nutzers abgefragt (Gedeon 2007, S. 31; Kinman 2009, S. 25). Gleichzeitig wird versucht durch Kombination von Nutzungsstatistiken mit Ergebnissen bibliometrischer Analysen wie z.B. Zitations- und Publikationshäufigkeit dem tatsächlichen Nutzen näher zu kommen (z.B. Kurtz et. al. 2005). Derzeit besteht der Konsens sicher darin, dass die Nutzungsstatistiken ein Indiz für den Nutzen einer elektronischen Quelle sind. Nicht mehr, aber auch nicht weniger.

Probleme, Fehler und Schwierigkeiten

Die Bearbeitung und die Interpretation von Statistiken elektronischer Ressourcen wird sowohl von definitorischen, technischen und systemimmanenten Mängeln als auch auf Grund von Lücken im derzeitigen COUNTER-Standard und der fehlenden systematischen Erfassung äußerer Einflussfaktoren behindert.

⁵ Diese Aussage beruht auf einer Diskussion im Rahmen der Session „Bibliotheken für die Forschung“ des 98. Bibliothekartags in Erfurt. Eine schriftliche Aufzeichnung existiert nicht.

Nutzungsstatistiken von lizenzierten Produkten werden (in den meisten Fällen) durch die Anbieter bereitgestellt und erfordern von den Bibliotheken blindes Vertrauen. Denn unklar ist, wie die berichteten Zahlen von den Anbietern erhoben werden. Dass dies keine exakten Angaben sein können, erläuterte z.B. Tobias Keppler im Rahmen der 3. Bremer E-Book-Tage (Kirchgäßner 2009). So können verschiedenen Nutzer, die über eine Proxyadresse zugreifen, nur als ein Nutzer gezählt werden, ein Download kann auf Grund von Zwischenspeichern nicht gezählt werden, etc.. Die Behauptung, dass nachvollziehbare und exakte Messungen von Downloads derzeit nicht möglich sind, wurde durch den Wiley-Verlag in einem informellen Gespräch bestätigt. Dies bedeutet, dass die ermittelten Statistiken derzeit nicht die realen Nutzungszahlen wiedergeben, wie auch Ferguson im Vergleich mit lokal gemessenen Daten zeigt⁶ (Ferguson/Rosemary 2005). Diese Problematik betrifft zwar sicher alle Anbieter, jedoch in unterschiedlichem Maße ausgeprägt. Die Vergleichbarkeit der Statistiken wird bereits an dieser Stelle gemindert. Augenscheinlich wird dies auch, wenn Nutzungszahlen mehrerer Monate plötzlich rückwirkend vom Verlage korrigiert werden, wie kürzlich vom Springer-Verlag im Bereich E-Books berichtet.

Die Beeinträchtigung durch technische Voraussetzungen wird verstärkt, wenn, wie in der Fraunhofer-Gesellschaft, durch eine Vielzahl von verschiedenen Instituten einzelne Statistiken erforderlich sind. Die Zuordnung zu den Statistiken erfolgt mittels IP-Adressen, welche sich nicht nur in regelmäßigen Abständen ändern, sondern sich bei geografischer Nähe auch mit z.B. den Adressräumen von Universitäten überschneiden können. Die Fehlergefahr durch falsche Zuordnung ist daher ständig gegeben und kann gleichzeitig nur in seltenen Fällen identifiziert werden. Solange von den Verlagen an dieser Stelle nicht systematisch offen und transparent gearbeitet wird, bleiben Nutzungsstatistiken eine Ver- oder vielmehr Misstrauensangelegenheit.

Vor Beginn jeder Analyse von Nutzungsstatistiken ist es notwendig diese auch lokal zu speichern. Was trivial klingt, entpuppt sich in der Praxis als komplexe Herausforderung. In der Fraunhofer-Gesellschaft existiert ein hand-breiter, penibel gepflegter Ordner, welche die Beschreibung des Weges zu den Nutzungsstatistiken aller lizenzierten Anbieter enthält. Darin werden Fragen beantwortet wie: Gibt es ein Portal oder wird die Statistik per E-Mail zugesandt? Und wenn per E-Mail – dann automatisch oder nur auf Nachfrage? Und wenn es ein Portal ist – wie sind die einzelnen Reiter und Links benannt, braucht man ein (oder mehrere) Passwort(e), welche

⁶ Wie zu erwarten war ist die von den Verlagen ermittelte Zahl wesentlich höher...

Formate werden angeboten und existiert eine Export-Funktion oder ist händisches kopieren notwendig etc.?? Die Heterogenität der Zugriffswege erfordert einen erheblichen Zeitaufwand, welcher dadurch alleine für die Bereitstellung der Statistiken aufgewendet werden muss (Whitehead 2007).

Diese Schwierigkeiten der Ermittlung und Bereitstellung machen jedoch insgesamt nur den kleineren Teil der Problematik im Zusammenhang von Nutzungsstatistiken aus.

Dies beginnt damit, dass, obwohl sich die Zahl der COUNTER-zertifizierten Anbieter seit 2005 mit jetzt 110 Anbietern⁷ fast verdoppelt hat, noch immer ein beachtlicher Teil vor allem von nationalen oder hoch-spezialisierten Verlagen und Datenbankproduzenten nicht dem COUNTER-Standard folgen. Zwar entsprechen die angebotenen Zahlen hier häufig den Standards – die Überprüfung dessen erfordert jedoch eine zeitaufwendige persönliche Kommunikation mit den Anbietern. Gleichzeitig führt die viel gelobte Einheitlichkeit durch COUNTER auch zu negativen Einschränkungen in den von den Anbietern angebotenen Auswertungen. So finden sich immer seltener zusätzliche und wichtige Informationen in Statistiken wie z.B. die thematische Kategorie in der eine Marktrecherche⁸ stattfand oder die Auswertung einzelner Suchschritte⁹.

Weiter bergen auch die zertifizierten Statistiken, COUNTER regelmäßig ignorierend, eine Vielzahl von größeren und kleineren Fehlern und Besonderheiten. Dies beruht zum einem auf noch immer bestehenden „Definitions-lücken“ im COUNTER-Glossar. So ist ein „Turnaway“ laut COUNTER ein nicht erfolgreicher Log-in begründet aus einer nicht erlaubten Anzahl von gleichzeitigen Nutzern¹⁰. Jedoch sind in dieser Zahl auch alle erfolglosen Log-Ins durch Systemprobleme oder aber fehlender Lizenz beinhaltet (Taylor-Roe/Spencer 2005, S. 129). Ein anderes Beispiel findet sich in der Anwendung der Definition von „Search“. Zwar beschreibt dies COUNTER als intellektuell ausgelöste Abfrage¹¹, doch wird von den Anbietern die sich ständig wiederholenden Abfragen eines einmal eingerichteten Alerts mitgezählt, da nach Ihrer Interpretation die einmalige Einrichtung einer bewussten (und damit intellektuellen)

⁷ Stand Mai 2010. URL: <http://www.projectcounter.org/r3/cv.pdf>

⁸ Frost & Sullivan

⁹ SciFinder (American Chemical Society)

¹⁰ COUNTER, 2010: Appendix A: Glossary of Terms, S. 12. http://www.projectcounter.org/r3/r3_A.doc

¹¹ COUNTER, 2010: Appendix A: Glossary of Terms, S. 8. URL: http://www.projectcounter.org/r3/r3_A.doc

Entscheidung entsprach. Dieses Vorgehen wäre vergleichbar damit, wenn Bibliotheken Verlängerungen ohne Unterscheidung als neue Ausleihe zählen würden...

Bei der eigenen Kontrolle der Statistiken, und auch in Berichten der Mailingliste lib-stats, finden sich wiederholt verschiedene Fehler: Zeilen fehlen, XML-Dateien werden falsch erstellt, Summen werden nicht korrekt addiert, Zahlen verändern sich ohne ersichtlichen Grund zwischen 2 Abrufen von ein und derselben Tabelle, die Trennzeichen einer CSV-Datei sind nicht einheitlich etc.. An dieser Stelle ist es nicht möglich eine Übersicht über alle bereits aufgetretenen Fehler zu berichten - es lässt sich aber vermuten, dass jeder Bibliothek oben genannte oder ähnliche COUNTER-Abweichungen begegnet sind.

Eine Herausforderung für Verlage und Bibliotheken stelle jedoch in jedem Fall die Unbenennung oder der Verkauf von Zeitschriften an andere Verlage da, welches es fast unmöglich macht, den Nutzungsverlauf von Zeitschriften kontinuierlich zu verfolgen. Die Änderung von ISSNs, die ungleiche Behandlung in der Zuordnung von Print- und Online-ISSNs und die unkommentierte Verwendung mal von Print-, mal von Online-ISSNs in den Statistiken von verschiedenen Verlagen führt zu Zuständen der Verzweiflung. Manche versuchen das Problem über den Aufbau einer Abgleich-Datenbank, andere durch die Programmierung von Computerprogrammen zu lösen. In der Mailingliste lib-stats wurde im Juli 2010 diskutiert, ISSNs durch Permanent Links oder ähnliche Formate zu ersetzen. Die Forderung nach einer dauerhaften Identifizierungsmöglichkeit ist logisch und wichtig, zeigt dabei jedoch auch das Scheitern der Konzepte „ISSN“ und „ISBN“.

Externer Einflussfaktoren auf Nutzungsstatistiken

Die Idee der Erstellung und Verwendung von Nutzungsstatistiken beruht auf einer Grundannahme: ein Nutzer findet durch Recherche, Empfehlung, Zitat oder ähnliches z.B. einen Volltext und lädt diesen auf seinen Computer, weil ihn das Thema, die Zeitschrift, der Titel, der Autor oder das Abstract interessiert. Eine höhere Anzahl von Downloads für eine Zeitschrift bedeutet also, dass diese für den Nutzerkreis interessanter und damit wichtiger ist...oder?

Verschiedene Studien haben gezeigt, dass neben Interesse noch weitere Einflussfaktoren eine Rolle spielen. Davis und Price zeigten 2006 in einem Vergleich von PDF- und HTML-Downloads von sechs COUNTER-zertifizierten Anbietern mit zum Teil gleichem Zeitschriftenangebot, dass die Höhe der Nutzung und das genutzte Format abhängig ist vom Design der Titelanzeige und der gewählten Verlinkung

(Davis/Price 2006). Die Zahl der Downloads hängt daher z.B. davon ab, ob zunächst eine bibliographische Titelanzeige erscheint, ob die Verlinkung auf die PDF-Version direkt gezeigt wird oder erst nach Aufruf der HTML-Version, ob die HTML-Version direkt erscheint und natürlich ob von Suchmaschinen wie z.B. Google auf den Volltext oder zunächst auf die bibliographische Angabe verlinkt wird. Welche Rolle die Verlinkung von Suchmaschinen spielt, zeigte sich bei IEEE in der ersten Hälfte des Jahres 2009. Hier war durch die Änderungen in der Linkgenerierung von IEEE die Verlinkung durch Google nicht intakt und die Downloads sanken weltweit¹² um ca. 50%.

In mehreren Studien konnte berichtet werden, dass die Höhe der Downloads auch in Abhängigkeit von der Präsentation bzw. dem Einsatz von LinkResolvern¹³ steht. Kinman berichtet gar von einem Anstieg von 50% durch die optimierte Einstellung des LinkResolvers (Kinman 2009). Auch McDonald fand eine signifikante Steigerung durch die Einführung eines LinkResolvers (McDonald 2007b). Darüber hinaus berichtete er von weiteren Einflussfaktoren auf die Höhe der Nutzung, wie die Anordnung der Quellen im Katalog, die Vorhaltung von lokalen Kopien für den wiederholten Gebrauch oder die Gestaltung der Zwischenseite des LinkResolvers SFX. Der Unterschied in der Nutzungshöhe zwischen Anbietern wird also nicht alleine von den Inhalten, sondern auch durch die Art der Verlinkung und die Präsentation der Angebote sowohl von Seiten des Anbieters als auch durch die Bibliothek bestimmt. Als eine Maßnahme zum Ausgleich dieser Unterschiede schlagen Davis und Price einen „adjustment factor“ vor, welche die Vor- und Nachteile einer Anbieter-Oberfläche bewertet und die Downloadzahlen verschiedener Anbieter dadurch vergleichbar macht (Davis/Price 2006, S. 1247).

Nutzungsstatistiken versus Lizenzvereinbarungen

Bei dem Versuch, belastbare Aussagen über den Erfolg der eigenen Bestandspolitik zu generieren, sind Nutzungsstatistiken das einzig derzeit objektive vermittelnde Instrument, wie einführend berichtet. Dabei wird in der Regel sowohl die Gesamtnutzung der Produkte eines Anbieters (vor allem im Vergleich zu anderen Anbietern) als auch die Betrachtung von einzelnen Produkten bzw. Titeln berücksichtigt.

¹² Dies ist eine Schätzung aus den Kommentaren der Mailingliste lib-stats, eigenen Beobachtungen und den Austausch mit anderen Organisationen.

¹³ Ein LinkResolver ist ein im Bereich digitaler Bibliotheken genutztes System zur kontextabhängigen Anzeige von Diensten oder Publikationen (Herb/Scholze 2007).

Bei der Bewertung der Gesamtnutzung der Produkte eines Anbieters gilt es jedoch zu berücksichtigen, dass in vielen Fällen die mit Volltext-Downloads ausgestatteten Zeitschriftentitel nicht mit den lizenzierten Titeln übereinstimmen. Zum kleineren Teil kommt es vor, dass Titel fehlen, häufiger noch kann beobachtet werden, dass für mehr Titel Nutzungsdaten vorliegen als von der Einrichtung lizenziert wurden. Dies kann seine Ursache in technischen Limitationen der Anbieter in Bezug auf die Freischaltung von Titeln haben, das Indiz für eine gezielte „Anfütterungs-Strategie“ oder ein Zeichen von Kulanz im Rahmen von Vertragsänderungen am Anfang eines Jahres sein. Aus der Gesamtsumme von Volltext-Downloads eines Zeitraums lässt sich also keine Aussage über die Nutzung eines lizenzierten Angebotes ableiten.

Verstärkt wird dieser Effekt in Deutschland durch das Lizenzierungskonzept der Nationallizenzen¹⁴. Produkte können sowohl durch verhandelte Lizenzen, als auch durch kostenfreie bzw. kostengünstige Nationallizenzen erworben werden. Eine Unterscheidung der verschiedenen Lizenzmodelle innerhalb der Nutzungsstatistiken ist auch mit großem Aufwand nicht vollständig möglich. Seit der Einführung von „COUNTER Release 3“ ist bei den meisten Anbietern eine Statistik über sog. Backfiles oder Archive erhältlich, welche sich in großen Teilen mit dem Zeitraum der Archiv-Nationallizenzen decken. Doch sind die Einteilung der Backfiles und der Zeiträume der Archiv-Lizenzen der Nationallizenzen meist um einige Jahre verschieden. In Diskussionen ist oft zu merken, dass diese Unterscheidung für den Großteil der Einrichtungen in Deutschland zu vernachlässigen ist. Doch, ausgehend vom momentanen Diskussionsstand könnte ein Szenario sein, dass in Zukunft auf Grund der auslaufenden Förderung durch die Deutsche Forschungsgemeinschaft (DFG) für die Nationallizenzen steigende Beträge zu investieren sind. Spätestens dann werden die Fragen entstehen: Wollen wir das Produkt? Brauchen wir das wirklich? Wie viel können wir verantworten dafür zu bezahlen? Und damit auch: nutzen wir das eigentlich? Die Frage wird in den wenigsten Fällen mit Hilfe von Statistiken exakt zu beantworten sein. Eine Möglichkeit bietet COUNTER derzeit bereits durch das Konzept der Statistik „Journal Report 5 (JR5)“. Hier werden die Volltext-Downloads pro Monat pro Zeitschrift pro Jahrgang angegeben. Jedoch sind die Anbieter lediglich verpflichtet entweder JR5 oder „Journal Report 1a“ anzubieten, welcher „nur“ die getrennte

¹⁴ Weitere Hintergründe siehe: www.nationallizenzen.org

Ausweisung von Archiven¹⁵ vorgibt¹⁶.

Die Aussagekraft von Nutzungsstatistiken in Bezug auf die Lizenzen ist zudem gänzlich abhängig von der genauen und aktuellen Kenntnis der eigenen Lizenzen. Insbesondere wenn eine Bibliothek bzw. Einrichtung Lizenzen nicht nur als einzelne Organisation, sondern auch konsortial bzw. „quasi-konsortial“, wie im Falle der Fraunhofer- oder auch der Max-Planck-Gesellschaft, erwirbt. Hier sind die Kenntnisse über lokal in den Instituten bzw. bei den Konsortialteilnehmern erworbenen Lizenzen nur unzureichend vorhanden, so dass z.B. in der Gesamtstatistik eine nicht-lizenzierte Zeitschrift eine Nutzung aufweist, welches letztendlich nur auf den Mitgliedszugang eines Mitarbeiters zurückzuführen sein kann. Kombiniert mit den ungenauen Abgrenzungen der Lizenzen durch die Verlage ist dann eine Aussage über die Gesamtzahlen nicht mehr möglich.

Auch der Blick auf die Nutzung einzelner Produkte bzw. Titel ist wenig erfolgversprechend. Selbst wenn die oben berichtete ISSN-Problematik vernachlässigt wird, bleibt in jedem Fall die Frage nach dem Sinn einer solchen Betrachtung. Denn in der Regel können Zeitschriftentitel oder E-Books nicht einzeln erworben bzw. abbestellt werden. Logische Konsequenz dieser Erkenntnis ist die Forderung an die Anbieter nach stärker differenzierbaren Vertragsmodellen und die Abkehr von Pakettlizenzierung (Best 2007, S. 206). Gedeon widersprach dieser Schlussfolgerung 2007 mit einer eindrucksvollen Studie, in welcher er belegte, dass die Kosten der Lizenzierung von Einzeltiteln durch den erhöhten Verwaltungsaufwand jegliche Einsparung durch mögliche Abbestellungen übertreffen (Gedeon 2007).

Fazit - „So, what should we do, throw up our hands and forget the whole thing?“

(Ferguson 2003, S. 94)

Also: Nutzungsstatistiken beruhen auf falschen technischen Voraussetzungen, sind extrem zeitaufwendig und fehlerhaft, berücksichtigen nicht den Kontext der Nutzung, entsprechen nicht dem lizenzierten Umfang und sagen eigentlich auch nichts über den tatsächlichen Nutzen einer elektronischen Ressource aus. Zusätzlich sind die Daten verschiedener Anbieter auf verschiedenen Ebenen grundsätzlich nicht

¹⁵ Archiv = Non-current collections of journals, books, articles, or other publications that are preserved because of their continuing value and which are frequently made available by publishers as separate acquisitions. COUNTER 2010: The Code of Practice Release 3, S. 5.
<http://www.projectcounter.org/r3/Release3D9.pdf>.

¹⁶ COUNTER 2010: The Code of Practice Release 3, S.13.
<http://www.projectcounter.org/r3/Release3D9.pdf>.

vergleichbar und die Abbestellung von Einzeltiteln weder möglich noch sinnvoll. Im Gegenzug verwenden die Anbieter diese Zahlen, um hohe Zeitschriftenpreise mit niedrigen „Cost per Download“-Zahlen“ zu begründen (Davis/Price 2006, S. 1243). Sind Nutzungsstatistiken am Ende nichts weiter als ein zeitaufwendiges Eigentor für die Bibliotheken und Informationseinrichtungen?

Ferguson kam 2003 zu dieser Erkenntnis und die University of Hong Kong beschloss „to pull back on our efforts to create homogenous, comparable statistics...“

(Ferguson 2003, S. 94). 2005 revidierte er diesen Beschluss mit der Einschränkung: „Usage statistics were not used for budget or staffing purposes but were largely used to help guide marketing the continued buildup of our digital collections.“ (Ferguson 2005, S. 4)

Unter den derzeitigen Gegebenheiten und unter Berücksichtigung aller Dimensionen von Nutzungsstatistiken, sind diese derzeit sicher nicht als Entscheidungsgrundlage für oder gegen die Lizenzierung von elektronischen Ressourcen geeignet. Jedoch zeigt eine hohe Download-Zahl möglicherweise eine Tendenz, welche unter Einbeziehung weiterer Faktoren wie Aufbau und Design der Oberfläche und Bekanntheit des Angebotes ein Hinweis auf einen möglichen Nutzen sein kann. Taylor-Roe und Spencer betonen in diesem Zusammenhang auch die Möglichkeit über die Nutzungszahlen mit den Kunden in Kontakt zu treten, so dass z.B. ermittelt werden kann: brauchen die Kunden ein bestimmtes Produkt nicht oder sind die Nutzungszahlen aus anderen Gründen niedrig? (Taylor-Roe/Spencer 2005, S. 129). Weiter führen die Autoren aus, dass Nutzungszahlen sich ideal für Benchmarking oder auch Rechtfertigung gegenüber Geldgebern eignet (Taylor-Roe/Spencer 2005, S. 126). Versehen mit dem Zusatz, dass die Nutzungszahlen nicht absolut, sondern eher als Tendenz zu verstehen sind, ist der Einsatz von Nutzungsstatistiken als Beleg für den Wert einer Informationseinrichtung oder Bibliothek ideal.

In verschiedenen Studien (McDonald 2007a, 2007b, Blečić/Fiscella/Wiberley Jr. 2007) werden Nutzungsstatistiken für die Beantwortung von gezielten Fragestellungen eingesetzt: Ändert sich die Nutzung durch verbessertes Marketing, Präsentation, die Installation einer Metasuche oder eines LinkResolvers? Auch finden Nutzungsstatistiken ihren Einsatz in der Analyse des Nutzerverhaltens – vor allem in Bezug auf die Veränderungen der Print-Nutzung gegenüber e-Journals (Rowlands 2007). Nutzungsstatistiken können damit Bibliotheken und Informationseinrichtungen dabei

unterstützen zwar nicht den Nutzerbedarf, jedoch das Nutzerverhalten besser zu verstehen.

Der Weg zu aussagekräftigen Nutzungsstatistiken, welche auch für Lizenzierungsentscheidungen eingesetzt werden können, ist noch lang – und so mancher Umweg wird notwendig sein. Bis dahin sollte die Arbeit mit Statistiken in Bibliotheken und Informationseinrichtungen zum selbstverständlichen Tagesgeschäft gehören. Nur so, ergänzt durch die Zusammenarbeit mit internationalen Projekten wie COUNTER und im ständigen Austausch untereinander als auch mit den Anbietern, kann es gelingen, Fehler und Schwierigkeiten zu verringern. Vielleicht ist es dann irgendwann auch über den Berg kürzer als zu Fuß...

Literatur

Baker, Gayle / Read, Eleanor J. (2008): Vender-supplied usage data for electronic resources: a survey of academic libraries. In: *Learned Publishing*, 21. Jg, S. 48-57.

Best, Rickey: Lies, Damn Lies, and Usage Statistics: What's a Librarian to Do? In: Fowler, David C.: *Usage Statistics of E-Serials*. Binghamton, S. 199-214, hier: S. 205 + S.209

Blecic, Deborah D. / Fiscella, Joan B. / Wiberley Jr., Stephen E. (2007): Measurement of Use of Electronic Resources: Advances in Use Statistics and Innovations in Resource Functionality. In: *College & Research Libraries*, 68. Jg., S. 26–44.

Bollen, Johan / Rodriguez, Marko A. / Van de Sompel, Herbert (2007): MESUR: usage-based metrics of scholarly impact. In: *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, Vancouver, BC, Canada, S. 474.

Davis, Philip M. / Price, Jason S. (2006): eJournal Interface Can Influence Usage Statistics: Implications for Libraries, Publishers, and Project COUNTER. In: *Journal of the American Society for Information Science and Technology*, 56. Jg, Heft 9, S. 1243-1248.

Ferguson, Anthony (2003): Back Talk – Use Statistics: Are they worth it? In: *Against the grain*, 14. Jg, Heft 6, S. 93-94. URL: http://www.against-the-grain.com/TOCFiles/Backtalk14_6.pdf. Stand: 22.07.2010

Ferguson, Anthony / Rosemary, Gayle (2005): Usage Statistics at Hong Kong University: From Fun to Fundamental in Just a Few Years. In: Borghuis, Marthyn: *What Counts and What Doesn't: An Insider's Guide to Usage Reports*, S.4-5. URL: <http://libraryconnect.elsevier.com/lcp/0701/LCP070101.html>. Stand: 22.07.2010

Gedeon, Randle (2007): Are we really balancing the Ledger with E-Journals? In: Fowler, David C.: *Usage Statistics of E-Serials*. Binghamton, S.25-42

Ulrich Herb, Ulrich / Scholze, Frank: Nutzungsstatistiken elektronischer Publikationen. In: *Zeitschrift für Bibliothekswesen und Bibliographie*. 2007, Heft 4-5, S. 234-237

Kinman, Virginia (2009): E-Metrics and Library Assessment in Action. In: *Journal of Electronic Resources Librarianship*, 21. Jg., s. 15-36.

Kirchgäßner, Adalbert (2009): Die Weiterentwicklung elektronischer Bücher. Bericht vom Dritten Bremer E-Book-Tag. In: Bibliothek aktuell, Nr. 89 (2009). URL: <http://w4.ub.uni-konstanz.de/ba/article/viewFile/3177/2979>. Stand: 21.07.2010

Kurtz, Michael J. et al. (2005): The Bibliometric Properties of Article Readership Information. IN: JASIST 56, S. 111–128.

McDonald, John (2007a): Understanding Journal Usage: A Statistical Analysis of Citation and Use. In: Journal of the American Society for Information Science and Technology, 58. Jg., Heft 1, S. 39-50

McDonald, John (2007b): Usage Statistics & Information Behaviors: Understanding User Behavior with Quantitative Indicators. Paper presented at Understanding the Data Around Us: Gathering and Analyzing Usage Daten, NISO Usage Data Forum, November 1-2, Dallas, Texas. URL: <http://niso.kavi.com/news/events/niso/past/usage07/usage07mcdonald.pdf>. Stand: 22.07.10

McDonald, John (Ed.) (2009): Journal of Electronic Resources Librarianship : Usage Statistics of Electronic Resources, 21. Jg, Heft 1

Moore, Will / MacCreery, Nancy / Marlow, Martin (2010): Usage Measurement for Digital Content. Prepared on behalf of Springer.
http://www.springer.com/cda/content/document/cda_downloaddocument/White+Paper+-+usage+measurements+for+digital+content?SGWID=0-0-45-952937-0. Stand: 21.07.2010

Peters, Thomas A. (2002): What's the use? The value of e-resource usage statistics. In: New Library World, Volume 103, S. 93-47.

Rowlands, Ian (2007): Electronic journals and user behavior: a review of recent research. In: Library & Information Science Research, 29. Jg, S. 369-396

Shepherd, Peter T. (2007): Final Report on the Investigation into the Feasibility of Developing and Implementing Journal Usage Factors.
<http://www.uksg.org.uk/sites/uksg.org/files/FinalReportUsageFactorProject.pdf>. Stand: 20.07.2010

Taylor-Roe, Jill / Spencer, Cliff: A librarian's view of usage metrics: through a glass darkly? In: Serials, 18. Jg., Heft 2, S. 124-131

Whitehead, Derek: Why have statistics? ... And what? Paper presented at the CEIRC Forum on Usage Statistics for E-Resources, in Sydney, Australia.
<http://www.caul.edu.au/datasets/usagestatistics2007whitehead.pdf>. Stand: 21.07.2010